Chapter 7: Noncancer Health Endpoints

7.1 Introduction

The toxicity assessment consists of two components: hazard identification and dose response evaluation. Hazard identification determines whether exposure to an agent can cause an increase in the incidence of a particular adverse health effect (e.g., birth defects) and whether the adverse health effect is likely to occur in humans. Dose response evaluation quantitatively evaluates the toxicity information and characterizes the relationship between the' dose of the contaminant administered or received and the incidence of adverse health effects in the exposed population.

Noncancer toxicity values (e.g., reference doses) are derived from the quantitative dose-response relationship data (USEPA, 1989), from both human and animal studies, although human epidemiological studies are rarely available. Therefore, animal studies are the primary basis for developing noncancer health endpoint toxicity values. These values are expressed as reference doses (RfDs) for oral exposures and reference concentrations (RfCs) for inhalation exposures (see Chapter 1 for more details on their derivation). Noncancer health endpoints are addressed in the toxicity assessment section of a baseline risk assessment. The Science Advisory Board (SAB) (USEPA, 1990a) recommended to EPA that "for assessment of non-cancer human health risks the Agency should try to establish a risk assessment framework consistent with that used for carcinogens." But, at present, there are only proposed guidelines for reproductive risk (USEPA, 1988a and 1988b) and final guidelines for developmental risks (USEPA, 1991). Guidelines for deriving other noncancer toxicity values do not exist. Guidelines for neurotoxicity are being developed.

7.2 Discussion of Noncancer Health Endpoints in Statutes, Regulations, and Guidelines

7.2.1 Statutory and Regulatory Noncancer Toxicity Assessment Framework

As with radionuclide, CERCLA, SARA, and the NCP (USEPA, 1990b) do not specifically address noncancer health risks, but the broad definitions of pollutant or contaminant do include noncancer health endpoints (see Section 6.2.1). In responding to comments on alternative toxicity values proposed by Potential Responsible Parties (PRPs), the NCP stated:

> EPA will, of course, consider such public comments submitted on toxicity. However, it is important to note that the Superfund risk assessment process typically relies heavily on existing toxicity information or profiles that EPA has developed on specific chemicals. EPA believes that the use of a consistent data base of toxicological information is important in achieving comparability among its risk assessments.

This data base is the Integrated Risk Information System (IRIS). The NCP also stated that:

> . . . where no toxicological information is available in EPA's data base, then EPA routinely considers other available information, including information provided by PRPs or other interested parties . . . . Key assumptions and uncertainties in both contaminant toxicity and human and environmental exposure estimates must be documented in the baseline risk assessment, as well as the sources and effects of uncertainties and assumptions on the risk assessment results. Exposure assumptions or other information, such as additional toxicity information, may be evaluated to determine whether the risks are likely to have been under- or overestimated.

On the confidence of a toxicity value in IRIS, the NCP pointed out that "a high confidence in a toxicity value reflects a consensus that the value is not likely to change." The corollary to this would be that a value with a low confidence would be more likely to change.

7.2.2 General Guidance for Noncancer Toxicity Assessment

As discussed in the introduction, toxicity assessment contains two components hazard identification and dose response evaluation. Deriving a toxicity value requires toxicological expertise; therefore, it is beyond the scope of the Superfund Public Health Evaluation Manual (SPHEM) (USEPA, 1986a) and Risk Assessment Guidance for Superfund (RAGS) (USEPA, 1989) to give guidance on this aspect of the toxicity assessment. However, these documents do give guidance on what to do with the toxicity value once it has been derived and presented in IRIS.

Superfund Public Health Evaluation Manual (SPHEM]

SPHEM (USEPA, 1986a) discusses acceptable intakes for subchronic exposures (AIS) and acceptable intakes for chronic exposures (AIC)." AICS and AISs are derived from methods similar to RfDs, but the methodology was not as strict as that for RfDs, and Health Effects Assessment (HEA) documents. These values are determined by identifying the no-observed-adverse-effect-level (NOAEL) in a study and dividing the NOAEL by several uncertainty factors. The uncertainty factors are as follows: 10 for extrapolation from animals to humans, 10 for different sensitivities with the human population, and 10 for the use of a subchronic NOAEL to derive a chronic (AIC) value. The AIS and AIC values for chemicals that produce the same toxic effects are utilized in the risk characterization step to derive the hazard index. Many of the AIS and AIC values for noncarcinogenic effects were supplied in Appendix A of SPHEM.

Risk Assessment Guidance for Superfund (RAGS). Volume 1: Human Health Evaluation Manual (Part A)

RAGS points out that the toxicity assessment has already been performed for many chemicals. Toxicity values and information can be obtained from IRIS, Health Effects Assessment Summary Tables (HEAST) (USEPA, 1994), EPA drinking water criteria documents, drinking water Health Advisory summaries, ambient water quality criteria documents, air quality criteria documents, ATSDR, and EPA's

undergone extensive peer review; thus, it will require an experienced toxicologist familiar with EPA toxicity assessments to present successfully any new data in the baseline risk assessment.

Toxicity values are expressed as reference doses (RfDs) for oral exposures, and reference concentrations (RfCs) for inhalation exposures. They are both derived using the same methodology. The terms acceptable daily intakes (ADIs), acceptable intakes for chronic exposures (AIC), and acceptable intake-s for subchronic exposures (AIS) have been superseded by RfDs. Chronic RfDs are defined as "an estimate (with uncertainty spanning perhaps an order of magnitude or greater) of a daily exposure level for the human population, including sensitive subpopulations, that is likely to be without an appreciable risk of deleterious effects during a lifetime" (USEPA, 1989). Chronic RfDs are used to evaluate the noncancer effects for exposure periods between seven years and a lifetime; subchronic RfDs (RfD_s) are used to evaluate the noncancer effects for exposure between two weeks and seven years; and developmental RfDs (RfD_{dt}), are used for a single exposure event. RfD_ss do not appear in IRIS because they have not gone through verification by an intra-agency workgroup. RfD_{dt}s are still under development and have not been placed in IRIS or HEAST. Inhalation RfDs (RfD_i) are derived in a similar manner as oral RfDs, but the dynamics of the respiratory system and its diversity across species, as well as the differences in the physiochemical properties of the chemicals, must also be considered.

In cases where there is a lack of human data, analysts use animal data along with professional judgement to determine a value for RfD. Professional judgement involves an assessment of the relevance and scientific quality of the experimental studies, including the use of the appropriate animal model. If data from several animal studies are evaluated, EPA first seeks to identify animal models that are likely to be most relevant to humans based on biological rationale (e.g., metabolic and pharmacokinetic data). If a relevant animal model is not available, then "as a matter of science policy, the study on the most sensitive species (the species showing a toxic effect at the lowest administered dose) is selected as the critical study for the basis of the RfD" (USEPA, 1989). EPA assumes that humans are as sensitive to contaminants as the most sensitive animal species.

The assessment must then identify the critical toxic effect. "The effect characterized by the lowest-observed-adverse-effect-level' (LOAEL) after dosimetric conversions is referred to as the critical toxic effect" (USEPA, 1989). The critical toxic effect can include any non-cancer toxic end point, which includes reproductive, developmental, and genetic effects. "After the critical study and toxic effect have been selected, EPA identifies the experimental exposure level representing the highest level tested at which no adverse effects (including the critical toxic effect) were demonstrated. This highest 'no-observed-adverse-effect-level,' (NOAEL) is the key datum obtained from the study of the dose-response relationship" (USEPA, 1989). The NOAEL of the critical toxic effect is chosen because it is assumed that "if the critical effect is prevented, then all toxic effects are prevented" (USEPA, 1989).

The RfD is calculated by dividing NOAEL (or LOAEL) for the critical toxic effect by uncertainty factors (UFs) and a modifying factor (MF). The appropriate NOAEL (or LOAEL) is divided by the product of all the applicable UFs and the MF (Barnes and Dourson, 1988). The uncertainty factors are 10 for extrapolation from animals to humans, 10 for different sensitivities within the human population, 10 for the use of a subchronic NOAEL, and 10 if a LOAEL is used to derive the chronic RfD value. The MF reflects a "qualitative professional assessment of additional uncertainties in the critical study and in the entire data base for the chemical not explicitly addressed by the preceding uncertainty factors"

(USEPA, 1989). This value ranges from 1 to 10, and the default value is 1. The use of the RfD in the risk characterization step is explained in Chapter 5, Chemical Mixtures.

Assessors should discuss the degree of confidence in the toxicity value(s) in the risk assessment. RAGS notes that "The degree of confidence ascribed to a toxicity value is a function of both the quality of the individual study from which it was derived and the completeness of the supporting data base." The degree of confidence assigned to a toxicity value involves the use of professional judgement. RAGS states that the following factors add to the confidence of the evidence that the contaminant poses a hazard to humans

- similar effects across species, strains, sex, and routes of exposure;
- clear evidence of a dose-response relationship;
- a plausible relationship among data on metabolism, postulated mechanism of action, and the effect of concern;
- similar toxicity exhibited by structurally related compounds; and
- some link between the chemical and evidence of the effect of concern in humans.

According to EPA, "High uncertainty (low confidence, low strength of evidence) indicates that the toxicity value might change if additional chronic toxicity data become available. Low uncertainty (high confidence) is an indication that a value is less likely to change as more data become available, because there is consistence y among the toxic responses observed in different species, sexes, study designs, or in dose-response relationships" (USEPA, 1989).

Because professional judgement plays a predominant role in selecting the critical study and in calculating the confidence level in the toxicity value, legitimate differences in this judgement may be negotiated. For example, assessors may debate which critical study should be used to derive the RfDs. Another issue is whether the most sensitive species was chosen over the more biological appropriate animal model. The confidence in the toxicity value, on the other hand, is related to the study design of the experiment. Were proper laboratory protocols followed? Does a new study have a better study design? For a new toxicity value to be accepted, the new study will need to have a better study design than the already accepted critical study. An experiment that does not produce definitive results will have a lower confidence. A new study that has a better study design and gives more definitive results will be more easily accepted than a poorly designed study that produces uncertain results. Risks from the old study and the new study must both be presented.

7.2.3 Guidelines for Reproductive Toxicity Assessments

There are two proposed guidelines for reproductive toxicity assessments, Proposed Guidelines for Assessing Female Reproductive Risk (USEPA, 1988a) and Proposed Guidelines for Assessing Male Reproductive Risk (USEPA, 1988b). Final reproductive risk assessment guidelines have not been

published. These two documents give guidelines on laboratory testing protocols, toxicity indices, end points, and their interpretation. Both guidelines point out that there are no mathematical models developed for the dose-response analysis, therefore, RfDs are derived as previously described, by applying uncertainty factors to a NOAEL or LOAEL.

The RfD is derived from one dose point, NOAEL or LOAEL, and therefore the current RfD approach does not take the shape or slope of the dose-response curve into consideration when deriving the RfD value. Two different chemicals may potentially produce similar NOAELs or LOAELs, but have significant differences in the slope of the dose-response curve. Deriving a RfD from the NOAEL or LOAEL would produce similar toxicity values for both chemicals. Deriving the RfD from the dose-response curve in this hypothetical situation could produce different toxicity values, and therefore risk estimates. Both guidelines address the issues of sexual behavior, fertility, and pregnancy outcomes.

Proposed Guidelines for Assessing Female Reproductive Risk

These guidelines state:

> Ideally, study designs for all reproductive toxicity evaluations should include a high dose that produce-s some indication of maternal or adult toxicity (i.e., a level which produces a statistically significant reduction in body weight, weight gain, or specific organ toxicity, but no more than 10% mortality), a low dose which demonstrates a no observed effect level (NOAEL) for adult and offspring effects, and at least one intermediate dose level. A concurrent control group treated with the vehicle used for agent administration should be included (USEPA, 1988a).

Protocols used for both female and male reproductive risk include single and multiple generation studies. EPA is currently evaluating the continuous breeding protocol. Guidelines on conducting tests have been published by the Office of Pesticide Programs (USEPA, 1982) and the Office of Toxic Substances (USEPA, 1985). In addition to considering test protocol, these agencies also consider the appropriate route of exposure, number of animals per dose group, range between doses, and the relevance of the test system to humans. These guidelines point out the need for well conducted studies.

> Reliance on the NOAEL places substantial importance on the power of the study to detect low-dose effects. Poorly designed studies employing insensitive measures may produce higher NOAELs than those from a well-designed and well-conducted study. Risk assessments based on such data may underestimate the actual human risk. Conversely, use of a NOAEL from a study in which there were wide intervals between doses may overestimate the actual risk (USEPA, 1988a).

Many of the testing protocols call for both the female and the male to be dosed with the contaminant If adverse effects are seen with this testing protocol, it is difficult to determine whether the effect was mediated through the female, male, or both. To determine a female-only mediated effect, treated females must be mated with untreated males.

The indices that can be used for assessing female reproductive toxicity areas follows: male mating index, male fertility index, female mating index, female fertility index, female fecundity index, parturition index, gestation index, live litter size, live birth index, viability index, lactation index, weaning index, and preweaning index. The guidelines point out that "the gestation index... may provide limited information, since litters with only one pup are counted the same as those with more than one pup," and "the live birth index varies proportionally with the number of stillborn in each litter. The viability of all offspring at birth cannot always be observed immediately after birth. Cannibalization of the pups is another phenomenon that occurs in rodents and can obscure the interpretation of this index (USEPA, 1988a)." Other indices include alterations in the onset of puberty, alterations in the reproductive cycle, oocyte toxicity, premature reproductive senescence, ovary weight and morphology, uterine weight and histology, and pituitary gland weight and histology.

Proposed Guidelines for Assessing Male Reproductive Risk

Many of the issues addressed in the female reproductive test protocols are also applicable to the male reproductive test protocols. An additional testing protocol used for male reproductive risk is the dominant lethal test. Duration of dosing is critical for male reproductive risk because "damage that is limited to spermatogonial stem cells will not appear in the caudia epididymis or in ejaculates for 8 to 14 weeks, depending on the test species" (USEPA, 1988 b). For humans, spermatogenesis is 12 weeks in duration, damaged sperm will appear in the ejaculate anywhere from 1 to 12 weeks depending on which stage was affected. The study design must consider this delayed effect. With some agents that bioaccumulate, the full impact on a given cell type could be delayed, as could the impact on functional end points such as fertility. The dosing regime is different for the highest dose in this guideline, "the highest dose should induce toxicity but not mortality" (USEPA, 1988b). No explanation is given as to why mortality is not desired for male reproductive tests, and up to 10% mortality is acceptable for female reproductive tests. Test guidelines specify the "use of 20 males and enough females to produce at least 20 pregnancies for each treatment group in each generation in the multigeneration reproduction test. However, 20 pregnancies are often achieved by mating two females per male and using less than 20 males per treatment group" (USEPA, 1988b). The guidelines point out that "in such cases, the statistical treatment of the data should be examined carefully," and "in assessing male reproductive toxicity, the male remains the unit of statistical analysis. Using the female as the statistical unit inflates the sample size and may distort data analysis and interpretation" (USEPA, 1988b). To determine a male-only mediated effect, treated males must be mated with untreated females.

Male reproductive toxicity endpoints routinely evaluated for risk include the following: body weight, organ weights (testes, epididymides, seminal vesicles, prostate, and pituitary gland), organ histopathology, mating ratio, pregnancy ratio, litter size, pre- and post-implantation loss, number of live/dead pups, sex ratios, malformations, birth and postnatal weights, and survival. Organ weights are not the most sensitive measure of toxicity. For example, "damage to the testes may often be detected as a weight change only at doses higher than those required to produce significant effects in other measures of gonadal status" and "studies may report accessory sex gland weights with or without expression of the secreted fluids. Weights without fluids show less variability than weights in the presence of the secreted fluid" (USEPA, 1988b). Histopathological examination of the tissues is considered a sensitive tool for evaluating potential reproductive risks. Proper fixation of the tissue is important because, the "use of formalin fixation combined with paraffin embedding of the testis results in artifacts in control as well as

treated tissue" (USEPA, 1988b). Histopathological evaluations will not detect all reproductive effects (for example, genetic damage to the germ cell, decreased sperm motility, and an increase in abnormal sperm forms). Fertility and pregnancy outcomes through breeding studies are required to detect these effects, although these tests are considered to be insensitive as measures of reproductive injury. One reason for this is that "in some strains of rats and mice, sperm production can be reduced by 90% without compromising fertility" (USEPA, 1988b). Human sperm production does not have the extra capacity that rodents exhibit, and therefore the fertility potential of humans is far less than that of test species. Other male reproductive endpoints evaluated include the following: sperm count (for human studies this is very important), sperm morphology, sperm motility, sperm-cervical mucus penetration, *in vitro* fertilization, hormone evaluation, reproductive organ biochemical markers, and sexual behavior. For sexual behavior, the site of semen deposition can be strongly influenced by the sexual preparedness of the female rat. Reduced preparedness in the female rat could result in lower fertility index, which might erroneously be attributed to a spermatotoxic effect.

For both female and male reproductive risk evaluations, interpretation of these indices in new studies. should be performed by an experienced reproductive toxicologist. For new studies to be considered by EPA, their study designs must be better than current studies, and the results should be more definitive. In other words, their needs to be less uncertainty associated with the new study than the currently accepted studies.

7,2.4 Guidelines for Developmental Toxicity Assessments

Two final guidelines have been published, Guidelines for the Health Assessment of Suspect Developmental Toxicants (USEPA, 1986b) and Guidelines for Developmental Toxicity Risk Assessment (USEPA, 199la). These guidelines only address developmental toxicity, but they are closely related to the reproductive toxicity guidelines because of the organ systems involved. These guidelines have indicated that animal studies appear to be good predictors of human developmental toxicity. These studies have also shown that humans appear to be the most sensitive species, sometimes by a factor of 100 or more. As with reproductive toxicity evaluations, the study design and interpretation of the results are the two areas of interest that will be negotiated between a currently accepted study and any new study that is to be presented by DOE. Evaluations of study design and interpretation of the results should be performed by an experienced developmental toxicologist because "a great deal of scientific judgement based on experience with developmental toxicity data and with principles of experimental design and statistical analysis may be required to adequately evaluate such data" (USEPA, 1986b). The uncertainty associated with a new study should be less than the currently accepted study before it can be expected that EPA will seriously consider the new study.

Guidelines for the Health Assessment of Suspect Developmental Toxicants

The study design for developmental tests is the same as the female reproductive toxicity test protocols (see Section 7.2.3). These guidelines added, "The route of exposure should be based on expected human exposure considerations, although data from other routes may sometimes be useful, especially if supported by pharmacokinetic information."

These guidelines discuss 51 maternal and developmental toxicity endpoints that are evaluated for development risk assessments. It is beyond the scope of this document to discuss all of these endpoints and their significance. The most sensitive statistical endpoint is change in fetal weight. The need for careful interpretation of results was stated:

> Although statistical analyses are important in determining the effects of a particular agent, the biological significance of data should not be overlooked. For example, with the number of end points that can be observed in developmental toxicity studies, a few statistically significant differences may occur by chance. On the other hand, apparent trends with dose may be biologically relevant even though statistical analyses do not indicate a significant effect (USEPA, 199la).

One area of concern that the public commented on in the guidelines was developmental effects produced only at maternally toxic doses. The guidelines stated that "current information is inadequate to assume that developmental effects at maternally toxic doses result only from the maternal toxicity; rather, when the lowest observed effect level is the same for the adult and developing organisms, it may simply indicate that both are sensitive to that dose level. Moreover, the maternal effects may be reversible while effects on the offspring may be permanent." Several known human developmental toxicants (e.g., smoking, alcohol) produce their effects at minimally toxic adult doses (USEPA, 1986b). This issue was addressed further in the 1991 guidelines.

Guidelines for Developmental Toxicity Risk Assessment

Revisions to the 1986 guidelines include combining the hazard identification and dose-response evaluation, and developing the RfD_{dt} and RfC_{dt} values. "Hazard identification/dose-response evaluation involves examining all available experimental animal and human data and the associated doses, routes, and timing and duration of exposures to determine if an agent causes developmental toxicity and/or maternal or paternal toxicity in that species and under what exposure conditions." The current use of NOAELs or LOAELs to derive RfDs does not take into account the variability of the data or the slope of the dose response curve. EPA is evaluating the use of the benchmark dose method to derive RfDs to remove the limitations in using NOAELs and LOAELs. The approaches taken in these guidelines are based on four assumptions

1. A contaminant that causes an adverse effect in experimental animals will potentially pose a threat to humans following the appropriate exposure.

2. The four manifestations of developmental toxicity (death, structural abnormalities, growth alteration and functional deficits) are all of concern. The guidelines state that "there is usually at least one experimental species that mimics the types of effects seen in humans, but in other species tested, the type of developmental perturbation may be different. Thus, a biologically significant increase in any of the four manifestations is considered indicative of an agent's potential for disrupting development and producing a developmental hazard."

3. "The types of developmental effects seen in animal studies are not necessarily the same as those that may be produced in humans."

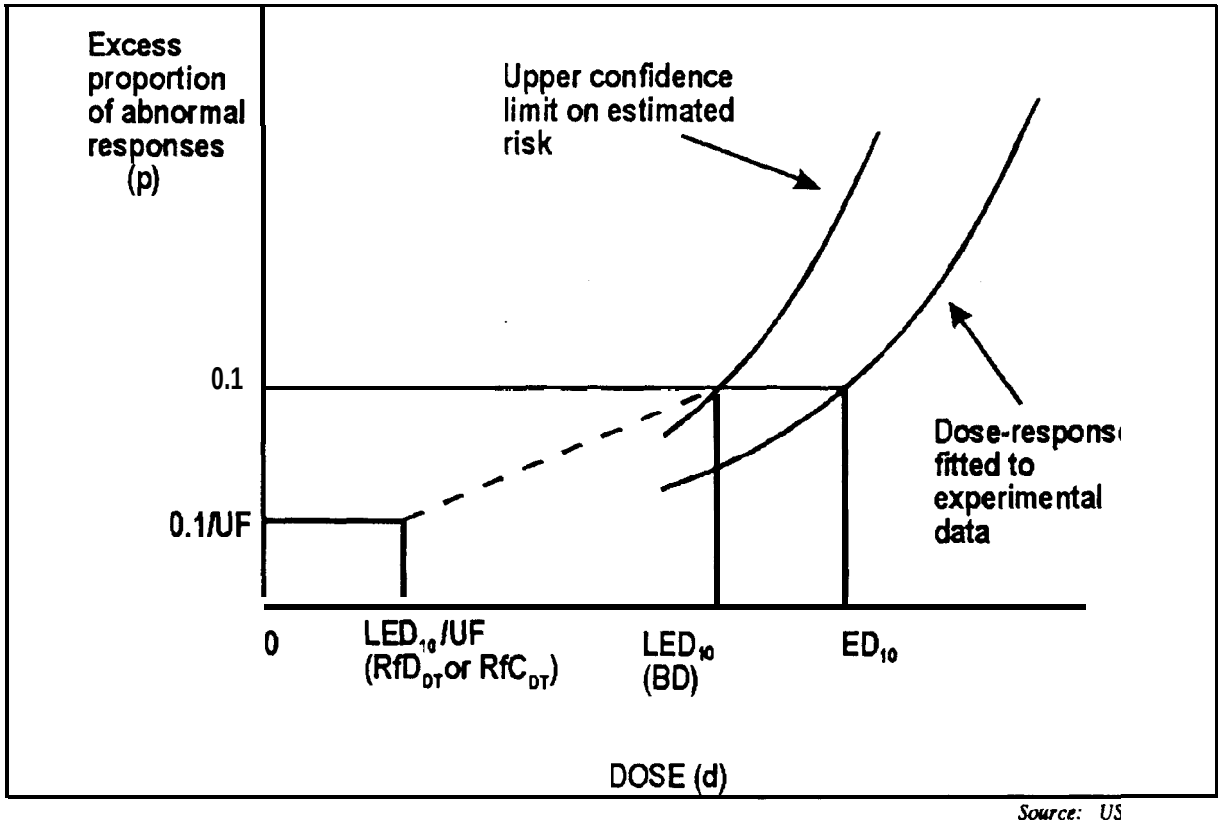
4. For developmental toxicants, a threshold is assumed for the dose-response curve.

The study design cited in the 1991 guidelines is not significantly different from that presented in the 1986 guidelines The only change was the use of a NOAEL for the low dose as opposed to a NOEL. It was explained that if the high dose produced excessive maternal toxicity (i.e., a dose that is significantly greater than the minimally toxic level), then results of the study would be difficult to interpret and of limited value. For selection of the test species, the SAB recommended that "the basic position of the Agency should be to use data from the most relevant species, and that use of data from the most sensitive species should be the default position" (USEPA, 1991). If a new, well-designed study is performed with a more relevant species, then the results from the new study can be a point for negotiation.

The end points of developmental toxicity considered in the 1991 guidelines are the same ones as in the 1986 guidelines, except that the 1991 guidelines added a discussion on functional deficits. Routine testing for functional deficits traditionally has not been required, but studies in developmental neurotoxicity are starting to be required by EPA under certain conditions. Generic developmental neurotoxicity test guidelines have been published (USEPA, 199lb). These guidelines also mention that offspring body weight is a sensitive indicator for developmental toxicity and they stress the importance of interpreting the results of a study from both a biological and statistical standpoint. They say that some variations may be statistically increased, but because they have a natural background of occurrence the variation may not be biologically significant (e.g., 12 or 13 ribs in rabbits). On the other hand, a variation may not be statistically significant but it may be biologically significant. These guidelines point out that "if a given study control group exhibits an unusually high or low incidence of postimplantation loss compared to historical controls, then scientific judgement must be used to determine the adequacy of the study for risk assessment purposes. "

Maternal toxicity and developmental toxicity were extensively discussed in the guidelines. It was again stated that if developmental effects are seen at minimal maternal toxic doses, then the contaminant is considered to be a potential developmental toxicant. When developmental effects are seen at minimal maternal toxicity levels, it is assumed that both the mother and the offspring are sensitive at the same dose. If maternal toxicity is produced at a lower dose than developmental effects, the results of the study become more uncertain. How this policy affects the risk assessment was stated as follows: "From a risk assessment point of view, whether a developmental effect is or is not secondary to maternal toxicity, does not impact on the selection of the NOAEL or other dose-response methodology." The reason for this is because the NOAEL is derived for the critical effect. As discussed earlier, it is assumed that if the critical toxic effect is prevented then all toxic effects will be prevented. For example, if maternal toxicity is observed at an exposure lower than developmental toxicity, the NOAEL will be derived from the maternal critical toxic effect, and it would be assumed that the developmental toxic effect would automatically be prevented. This issue requires careful evaluation by an experienced toxicologist to determine the exposure levels for maternal and developmental toxicity. Another associated issue that should be evaluated is a comparison of maternal toxicity to other adult toxicity values. The pregnant or lactating female may be more sensitive than the rest of the adult population,

Because the use of the NOAEL or LOAEL does not consider variations in the data or the shape of the dose-response curve, EPA has begun evaluating the benchmark dose. Figure 7.1 explains the benchmark dose approach. The SAB strongly suggested "the use of a benchmark dose approach to replace the NOAEL" (USEPA, 1991a). EPA is researching the mathematical models to be used in this approach.



This graphical illustration of the benchmark dose is based on Crump (1984) and Kimmel and Gaylor (1988). The benchmark dose (BD) is derived by modeling the data in the observed range, selecting an incidence level within or near the observed range (e.g., the effective dose to produce a 10% increased incidence of response, the ED_{10}, and determining the upper confidence limit on the model. The upper confidence value corresponding to, for example, a 10% excess in response is used to derive the BD which is the lower confidence limit on dose for that level of excess response, in this case, the LED_{10}. The RfD_{dt} or RfC_{dt} estimated by applying uncertainty factors (UF) to the BD would be greater than or equal to the BD/UF.

Figure 7.1 Benchmark Dose Approach

7.3 Issues and Regulator Dialogue

7.3.1 Noncancer Health Endpoint Issues

The statutory and regulatory history shows that noncancer toxic effects are to be included in the baseline risk assessment process. The history also indicates that alternative toxicity values from PRPs may

be considered, but that EPA prefers to use values that it has derived. Values supplied by PRPs are most Likely to be considered when there are no existing values or when the confidence in the existing value is low. Alternative toxicity values also can be presented in the uncertainty section to "determine whether the risks are likely to have been underestimated or overestimated" (USEPA, 1990b).

The guidance documents focused on the confidence placed in a toxicity value and the interpretation of toxicological studies. Toxicity values developed from studies with high uncertainty might be changed with new data. Professional judgement is involved in selecting the critical study, interpreting the results, and determining the level of confidence in a study. All of these judgments should be performed by an experienced toxicologist. There are no general guidelines for assessing noncancer toxicity effects. Proposed guidelines exist for reproductive risk, and final guidelines exist for developmental toxicity. EPA is beginning to research alternative methods to develop RfDs because the existing methodology does not take into consideration variations in the data or the shape of the dose-response curve.

Alternative Toxicity Values Can Be Presented in the Risk Assessment

Both the regulations and the guidelines indicate that alternative toxicity values may be considered by EPA, There are generally three reasons to present new toxicity values:

- If there are no existing value(s) in IRIS.
- If the value in IRIS is assigned a low confidence.
- If the estimation of risk using existing values in IRIS have been under-or overestimated.

It will require an experienced toxicologist, however, to assist DOE in negotiating new toxicity values. Toxicity values in IRIS are peer reviewed and are placed in the data base only after they have been accepted by a consensus of scientific opinion. Presentation of any new toxicity value will be working against a peer-reviewed, EPA-accepted value, if a value already exists in IRIS. The remaining issue points will discuss various aspects of studies that should be considered when presenting new toxicity values.

Appropriate Study Design

Study design includes selection of the appropriate animal model, and the dosing regime. EPA's policy is that the most relevant animal model, based on biological rationale, be selected first. "EPA first seeks to identify the animal model that is most relevant to humans based on a defensible biological rationale, for instance, using comparative metabolic and pharmacokinetic data" (USEPA, 1989). If a relevant model is not available, then the most sensitive species should be chosen, since it is assumed that humans are as sensitive as the most sensitive animal tested. If a new study is available that has used a more relevant animal model than the existing study, then DOE can negotiate the use of the new study. There must be biological and pharmacokinetic data available to prove that the new animal model is more relevant than the currently accepted test animal.

Dosing regime includes the concentration of the contaminant, schedule, and route of administration. Typically three doses are utilized in animal studies: 1) a high dose that produces toxicity, but not more than 10% mortality; 2) a NOAEL dose, and 3) an intermediate dose. Scheduling of the dose is critical for some toxicity endpoints (e.g., reproductive, developmental). For example, a contaminant may adversely affect only one stage of development or one stage of spermatogenesis. The route of administration should be relevant to human exposures, "the route of exposure should be based on expected human exposure considerations" (USEPA, 1986b). Any new study that is presented to EPA must follow the appropriate dosing regime before the study can be expected to be seriously considered by the Agency.

Selection and Interpretation of Toxicity Endpoints

The appropriate selection of toxicity endpoints for the NOAEL is very important because some endpoints are more sensitive than others, and some endpoints can give ambiguous results. For example, for female reproductive toxicity the gestation index "may provide limited information; since litters with only one pup are counted the same as those with more than one pup" (USEPA, 1988a). Some toxicity endpoints are time sensitive, therefore selection of when to measure the endpoint may be crucial for detecting a toxic effect. For example, contaminants may only effect one stage of spermatogenesis.

It is beyond the scope of this document to discuss all of the possible issues involved with the appropriate selection and interpretation of the various toxic endpoints; there are 51 endpoints for developmental toxicity, This chapter has discussed some of the issues for reproductive and developmental toxicity risks. An experienced toxicologist will be needed to assist in negotiating new studies.

7.3.2 Regulator Dialogue

All of the above issues must be considered when presenting any new studies to EPA. "The degree of confidence ascribed to a toxicity value is a function of both the quality of the individual study from which it was derived and the completeness of the supporting data base" (USEPA, 1989). The confidence in any study is associated with all of the above issues. Any new study that is presented to EPA should have a better study design and clearer interpretation of the results than the study used to derive the toxicity value in IRIS. If these conditions are not met, it cannot be expected that EPA will seriously consider the new study as an alternative to the existing study. Many of these issues can also be applied to cancer toxicity studies.

The above discussion indicates that EPA may consider alternative toxicity values in the risk assessment. It also indicates that alternative values will probably only be seriously considered when either the IRIS data base does not have a value, or when a value in the data base has a low confidence assigned to it. Alternative toxicity values may also be presented in the uncertainty section of the baseline risk assessment if it is believed that the risks are underestimated or overestimated when calculated using the values from IRIS. Neither CERCLA, SARA, nor the NCP gave guidance on how to conduct a toxicity assessment for noncarcinogens. Guidelines for deriving reproductive and developmental toxicity values are addressed in guidance documents. There are no guidelines for deriving other noncancer toxicity values.

7.4 References

Barnes, D.G. and Dourson, M. 1988. Reference Dose (RfD): Description and Use in Health Risk Assessments. Regulatory Toxicology and Pharmacology. Vol. 8, pp. 471486.

USEPA. 1982. Pesticide Assessment Guidelines, Subdivision F. Hazard Evaluation: Human and Domestic Animals. U.S. Environmental Protection Agency, Office of Pesticides and Toxic Substances, Washington, D.C. EPA-540/9-82-025.

USEPA. 1985. Toxic Substances Control Act Test Guidelines: Final Rules. U.S. Environmental Protection Agency, Office of Pesticides and Toxic Substances, Washington, D.C. 50 FR 39426 and 50 FR 39433.

USEPA. 1986a. Superfund Public Health Evaluation Manual (SPHEM). U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. Washington, D.C. EPA 540/1-86/060.

USEPA. 1986b. Guidelines for the Health Assessment of Suspect Developmental Toxicants. U.S. Environmental Protection Agency, Office of Health and Environmental Assessment. Washington, D.C. 51 FR 34028.

USEPA. 1988a. Proposed Guidelines for Assessing Female Reproductive Risk U.S. Environmental Protection Agency, Office of Health and Environmental Assessment. Washington, D.C. 53 FR 24834.

USEPA. 1988b. Proposed Guidelines for Assessing Male Reproductive Risk U.S. Environmental Protection Agency, Office of Health and Environmental Assessment. Washington, D.C. 53 FR 24850.

USEPA. 1989. Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual (Part A) (RAGS). U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. Washington, D.C. EPA 540/1-89/002.

USEPA. 1990a. Reducing Risk: Setting priorities and Strategies for Environmental Protection. U.S. Environmental Protection Agency, Science Advisory Board. Washington, D.C. SAB-EC-90-021.

USEPA. 1990b. National Oil and Hazardous Substances Pollution Contingent y Plan (NCP). U.S. Environmental Protection Agency. 55 FR 8666.

USEPA. 1991a. Guidelines for Developmental Toxicity Risk Assessment. Final. U.S. Environmental Protection Agency, 0ffice of Health and Environmental Assessment, Washington, D.C. 56 FR 63798.

USEPA. 199lb. Pesticide Assessment Guidelines, Subdivision F. Hazard Evaluation: Human and Domestic Animals. Addendum 10: Neurotoxicity, Series 81, 82, and 83. U.S. Environmental Protection Agency, Office of Pesticides and Toxic Substances. Washington, D.C. EPA 540/09-91-123.

USEPA. 1994. Health Effects Assessment Summary Tables (HEAST) Annual FY 1992. U.S. Environmental Protection Agency, Office of Emergency and Remedial Response. Washington, D.C. OERR 9200.6-303 (92-l); NTIS No. PB92-921199.